DINA: Open Source and Open Services - A modern approach for natural history collection management systems and research

Falko Glöckler‡, James Macklin§, Fredrik Ronquist|, Jana Hoffmann‡

‡ Museum für Naturkunde Berlin, Leibniz Institute for Evolution and Biodiversity Science, Berlin, Germany
§ Agriculture and Agri-Food Canada, Ottawa, Canada
| Swedish Museum of Natural History, Stockholm, Sweden

Corresponding author: Falko Glöckler (falko.gloeckler@mfn.berlin)

Abstract

The DINA Consortium ("**DI**gital information system for **NA**tural history data", https://dina-project.net) was formed in order to provide a framework for like-minded large natural history collection-holding institutions to collaborate through a distributed Open Source development model to produce a flexible and sustainable collection management system. Target collections include zoological, botanical, mycological, geological and paleontological collections, living collections, biodiversity inventories, observation records, and molecular data. DINA is funded by the participating member institutions. DINA Core Members are organizations or individuals who commit at least one half-time equivalent of resources to the development of the consortium goals, at least half of which should be available for code development.

The DINA system is architected as a loosely-coupled set of several web-based modules. The conceptual basis for this modular ecosystem is a compilation of comprehensive guidelines for Web application programming interfaces (APIs) to guarantee the interoperability of its components. Thus, all DINA components can be modified or even replaced by other components without crashing the rest of the system as long as they are DINA compliant. Furthermore, the modularity enables the institutions to host only the

components they need. DINA focuses on an Open Source software philosophy and on community-driven open development, so the contributors share their development resources and expertise outside of their own institutions.

One of the overarching reasons to develop a new collection management system is the need to better model complex relationships between collection objects (typically specimens), research data and associated workflows. We will present the enhancements provided by the approach of the DINA system focussing on the flexibility to plug in compliant components and accommodate additional (meta-)data and specimen related research data with the help of a generic data module.

Furthermore, we will discuss challenges in the governance of the development activities such as organizing the distributed code development of the core modules, the code review process and the choice of the software stack. These organizational challenges will be overcome with the help of a revised Memorandum of Understanding.

Keywords

natural history collection management, information system, open source, software development

Presenting author

Falko Glöckler

Presented at

Biodiversity_Next 2019